



UNITED STATES PATENT AND TRADEMARK OFFICE

ML

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|-------------|----------------------|---------------------|------------------|
| 09/976,098 | 10/15/2001 | Hiroaki Yoshino | 35.G2919 | 9468 |
| 5514 | 7590 | 07/12/2007 | EXAMINER | |
| FITZPATRICK CELLA HARPER & SCINTO 30 ROCKEFELLER PLAZA NEW YORK, NY 10112 | | | JACKSON, JAKIEDA R | |
| ART UNIT | | PAPER NUMBER | | |
| | | 2626 | | |
| MAIL DATE | | DELIVERY MODE | | |
| 07/12/2007 | | PAPER | | |

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

| | | | |
|------------------------------|------------------------|---------------------|--|
| Office Action Summary | Application No. | Applicant(s) | |
| | 09/976,098 | YOSHINO ET AL. | |
| | Examiner | Art Unit | |
| | Jakieda R. Jackson | 2626 | |

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).

Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) Responsive to communication(s) filed on 04 April 2007.
- 2a) This action is FINAL. 2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) Claim(s) 1,5-8 and 12-18 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) Claim(s) _____ is/are allowed.
- 6) Claim(s) 1,5-8 and 12-18 is/are rejected.
- 7) Claim(s) _____ is/are objected to.
- 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers'

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

| | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

Response to Amendment

1. In response to the Office Action mailed January 4, 2007, applicant submitted an amendment filed on April 4, 2007, in which the applicant amended and requested reconsideration with respect to independent claims 1, 8 and 15-18.

Response to Arguments

2. Applicant argues that Keiller involves comparing two user utterances, but does not suggest a comparison or matching rate between a user's utterance and a recording character string. In particular, Keiller discusses that "word model" are used for comparison. These "word models" are generated from a sequence of feature vectors that are output by a feature extraction routine, which extracts nine cepstral coefficients and one energy coefficient for each frame of input speech. Keiller, column 14, lines 14-29. In other words, Keiller teaches a 10-dimentional acoustic feature parameter. The method discloses by Keiller is, therefore, not disclosed as being capable of generating a character string. In response to applicant's argument that Keiller teaches a 10-dimentional acoustic feature parameter, the fact that applicant has recognized another advantage which would flow naturally from following the suggestion of the prior art cannot be the basis for patentability when the differences would otherwise be obvious. See *Ex parte Obiaya*, 227 USPQ 58, 60 (Bd. Pat. App. & Inter. 1985). Also, as pointed out in the office action dated April 4, 2007, Keiller teaches determination means for comparing a pattern of the recognized character string with a pattern if the recording character string stored in said storage means so as to obtain a matching rate

Art Unit: 2626

therebetween, and determining whether said matching rate exceeds a predetermined level (system checks whether training examples are consistent (column 15, lines 28-30) by computing the consistency scores (column 15, lines 53-65) and comparing the result again against the threshold (95%, column 16, lines 6-8). Therefore, Applicant's arguments are not persuasive.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

4. **Claims 1, 8, 15-18** are rejected under 35 U.S.C. 103(a) as being unpatentable over Keiller (USPN 6,560,575) in view of Jochumson (USPN 6,865,536) in view of Kirby et al. (USPN 6,226,615), hereinafter referenced as Kirby and in further view of Brown et al. (USPN 6,061,654), hereinafter referenced as Brown.

Regarding claims **1, 8, and 15**, Keiller discloses an apparatus, method and system for recording speech, to be used as learning data for recognizing input speech, comprising:

storage means for storing a recording character string indicating a sentence to be recorded (column 16, lines 12-19);

Art Unit: 2626

recognition means for recognizing input speech of the displayed sentence that a user reads out, and for obtaining a recognized character string (input is taken as two training examples: one a new example and one an already existing example; column 15, lines 25-35) corresponding to the stored recording character string pattern (column 16, lines 16-19);

determination means for comparing a pattern of the recognized character string with a pattern of the recording character string stored in said storage means so as to obtain a matching rate therebetween, and determining whether said matching rate exceeds a predetermined level (system checks whether training examples are consistent (column 15, lines 28-30) by computing the consistency scores (column 15, lines 53-65) and comparing the result again against the threshold (95%, column 16, lines 6-8); and

recording means for recording the input speech as the learning data for recognizing speech when it is determined by said determination means that said matching rate exceeds a predetermined level (if the results are consistent, they are used to generate a model for word being trained (column 15, lines 27-30), so inherently, the generated model is stored (recorded) to some memory means (see also column 16, lines 12-15), but does not specifically teach display control means, re-input instruction means and presentation means.

Jochumson discloses a speech correction device further comprising presentation means for presenting an unmatched portion between the recognized character string

Art Unit: 2626

pattern (what user has actually verbalized) and the recording character string pattern (what is expected; column 2, lines 53-65), to provide results or feedback.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Keiller's apparatus and method further comprising presentation means for presenting an unmatched portion between the recognized character string pattern and the recording character string pattern, as taught by Jochumson, to provides results and feedback to the user on how correct they were in stating the proper word or phrase (column 2, lines 53-65).

Keiller in view of Jochumson teaches storage means, determination means, recording means and presentation means, but does not specifically teach display control means and recognition means.

Kirby discloses a speech recognition device comprising a display control means for controlling displaying of the recording character string indicating the sentence to be recorded (prompting system to identify the words to be spoken that are presented; column 2, lines 31-39 with column 3, lines 1-17), to determine a new match between text and speech.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Keiller in view of Jochumson's apparatus and method wherein it comprises a display control means, as taught by Kirby, to determine a new match between text and speech, in order to try and regain synchronization (column 3, lines 51-65).

Art Unit: 2626

Keiller in view of Jochumson and Kirby teaches a storage means, display control means determination means, recording means and presentation means, but does not specifically teach a re-input instruction means.

Brown teaches a speech synthesis apparatus comprising a re-input means for issuing an instruction to input speech once again when it is determined by said determination means that the matching rate does not exceed the predetermined level (indicates that no such match exists, re-prompt the user to speak again; column 3, lines 28-52), to present the highest correct character string.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Keiller in view of Jochumson and Kirby's apparatus and method wherein it comprises a re-input instruction means, as taught by Brown, to present to the user with a positive match (column 3, lines 28-52).

Regarding **claim 16**, Keiller discloses a speech recognition method comprising:
a learning recognition step of recognizing input speech, of the displayed sentence that a user reads out, and for obtaining a recognized character string (input is taken as two training examples: one a new example and one an already existing example; column 15, lines 25-35);

a determination step of comparing a pattern of the recognized character string with a pattern of a recording character string indicating a sentence to be recorded so as to obtain a matching rate therebetween, and of determining whether said matching rate exceeds a predetermined level (system checks whether training examples are

Art Unit: 2626

consistent (column 15, lines 28-30) by computing consistency scored (column 15, lines 53-65) and comparing the result against a threshold (95%, column 16, lines 6-8));

a recording step of recording the input speech as the learning data for recognizing speech when it is determined in said determination step that said matching rate exceeds a predetermined level (if results are consistent, they are used to generate a model for word being trained (column 15, lines 27-30), so inherently, the generated model is stored (recorded) to a memory means (column 16, lines 12-19));

a learning step of performing learning on a speech model by using the input speech recorded in said recording step (the process described above provides general training of the model; column 16, lines 14-20); and

a recognition step of recognizing unknown input speech by using the speech model learned in said learning step (training data is used in general recognition; column 16, lines 14-20), but does not specifically teach display control means, re-input instruction means and presentation means.

Jochumson discloses a speech correction device further comprising presentation means for presenting an unmatched portion between the recognized character string pattern (what user has actually verbalized) and the recording character string pattern (what is expected; column 2, lines 53-65), to provide results or feedback.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Keiller's apparatus and method further comprising presentation means for presenting an unmatched portion between the recognized character string pattern and the recording character string pattern, as taught

Art Unit: 2626

by Jochumson, to provides results and feedback to the user on how correct they were in stating the proper word or phrase (column 2, lines 53-65).

Keiller in view of Jochumson teaches storage means, determination means, recording means and presentation means, but does not specifically teach display control means and recognition means.

Kirby discloses a speech recognition device comprising a display control means for controlling displaying of the recording character string indicating the sentence to be recorded (prompting system to identify the words to be spoken that are presented; column 2, lines 31-39 with column 3, lines 1-17), to determine a new match between text and speech.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Keiller in view of Jochumson's apparatus and method wherein it comprises a display control means, as taught by Kirby, to determine a new match between text and speech, in order to try and regain synchronization (column 3, lines 51-65).

Keiller in view of Jochumson and Kirby teaches a storage means, display control means determination means, recording means and presentation means, but does not specifically teach a re-input instruction means.

Brown teaches a speech synthesis apparatus comprising a re-input means for issuing an instruction to input speech once again when it is determined by said determination means that the matching rate does not exceed the predetermined level

Art Unit: 2626

(indicates that no such match exists, re-prompt the user to speak again; column 3, lines 28-52), to present the highest correct character string.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Keiller in view of Jochumson and Kirby's apparatus and method wherein it comprises a re-input instruction means, as taught by Brown, to present to the user with a positive match (column 3, lines 28-52).

Regarding claims 17 and 18, Keiller discloses a control program having computer readable program code and a speech recognition method, comprising:

a second program code unit for recognizing input speech of the displayed sentence that a user reads out, and for obtaining a recognized character string (input is taken as two training examples: one a new example and one an already existing example; column 15, lines 25-35);

a third program code unit for comparing a pattern of the recognized character string with a pattern of the recording character string so as to obtain a matching rate therebetween, and for determining whether said matching rate exceeds a predetermined level system checks whether training examples are consistent (column 15, lines 28-30) by computing consistency scored (column 15, lines 53-65) and comparing the result against a threshold (95%, column 16, lines 6-8);

a fourth program code unit for recording the input speech as the learning data for recognizing speech when it is determined by said determination step that said matching rate exceeds a predetermined level (if results are consistent, they are used to generate

Art Unit: 2626

a model for word being trained (column 15, lines 27-30), so inherently, the generated model is stored (recorded) to a memory means (column 16, lines 12-19);

a fourth program code unit for performing learning on a speech model by using the input speech recorded in said record step (the process described above provides general training of the model; column 16, lines 14-20); and

a eighth program code unit for recognizing unknown input speech by using the speech model learned in said learning step (training data is used in general recognition; column 16, lines 14-20), but does not specifically teach display control means, re-input instruction means and presentation means.

Jochumson discloses a speech correction device further comprising presentation means for presenting an unmatched portion between the recognized character string pattern (what user has actually verbalized) and the recording character string pattern (what is expected; column 2, lines 53-65), to provide results or feedback.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Keiller's apparatus and method further comprising presentation means for presenting an unmatched portion between the recognized character string pattern and the recording character string pattern, as taught by Jochumson, to provides results and feedback to the user on how correct they were in stating the proper word or phrase (column 2, lines 53-65).

Keiller in view of Jochumson teaches storage means, determination means, recording means and presentation means, but does not specifically teach display control means and recognition means.

Kirby discloses a speech recognition device comprising a display control means for controlling displaying of the recording character string indicating the sentence to be recorded (prompting system to identify the words to be spoken that are presented; column 2, lines 31-39 with column 3, lines 1-17), to determine a new match between text and speech.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Keiller in view of Jochumson's apparatus and method wherein it comprises a display control means, as taught by Kirby, to determine a new match between text and speech, in order to try and regain synchronization (column 3, lines 51-65).

Keiller in view of Jochumson and Kirby teaches a storage means, display control means determination means, recording means and presentation means, but does not specifically teach a re-input instruction means.

Brown teaches a speech synthesis apparatus comprising a re-input means for issuing an instruction to input speech once again when it is determined by said determination means that the matching rate does not exceed the predetermined level (indicates that no such match exists, re-prompt the user to speak again; column 3, lines 28-52), to present the highest correct character string.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Keiller in view of Jochumson and Kirby's apparatus and method wherein it comprises a re-input instruction means, as taught by Brown, to present to the user with a positive match (column 3, lines 28-52).

Art Unit: 2626

5. **Claims 5 and 12** are rejected under 35 U.S.C. 103(a) as being unpatentable over Keiller in view of Jochumson, Kirby and Brown and in further view of Crepy et al. (USPN 6,622,121), hereinafter referenced as Crepy.

Regarding **claims 5 and 12**, Keiller in view of Jochumson, Kirby and Brown discloses an apparatus and method for recording speech, to be used as learning data in speech recognition processing, but lacks wherein said presentation means presents the unmatched portion so as to identify the type of error as an insertion error, a deletion error, or a substitution error, as determined by said determination means.

Crepy discloses a speech correction device wherein said presentation means presents the unmatched portion so as to identify the type of error as an insertion error (insertions), a deletion error (deletions), or a substitution error (substitutions), as determined by said determination means (column 4, line 65 – column 5, line 11), to generate an error report.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Keiller in combination with Jochumson, Kirby and Brown apparatus and method wherein said presentation means presents the unmatched portion so as to identify the type of error as an insertion error, a missing error, or a substitute error, as taught by Crepy, to generate an error report from which various measurements may be derived (column 4, line 65 – column 5, line 11).

Art Unit: 2626

6. **Claims 6-7 and 13-14** are rejected under 35 U.S.C. 103(a) as being unpatentable over Keiller in view of Jochumson, Kirby and Brown and in further view of Baker (USPN 6,122,613).

Regarding **claims 6 and 13**, Keiller in view of Jochumson, Kirby and Brown discloses an apparatus and method for recording speech, to be used as learning data in speech recognition processing, but lacks wherein said presentation means simultaneously displays the recognized character string and the recording character string on a screen by changing a character attribute or a background attribute of an unmatched portion or a matched portion of at least one of the recognized character string and the recording character string.

Brown does not specifically teach a speech correction device wherein said presentation means simultaneously displays the recognized character string and the recording character string on a screen by changing a character attribute or a background attribute of an unmatched portion or a matched portion of at least one of the recognized character string and the recording character string (highlight uncertainty using reverse contrast; column 7, lines 1-16 and column 11, lines 23-30), to identify the error.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Keiller in view of Jochumson, Kirby and Brown apparatus and method wherein said presentation means simultaneously displays the recognized character string and the recording character string on a screen by changing

Art Unit: 2626

a character attribute or a background attribute of an unmatched portion or a matched portion of at least one of the recognized character string and the recording character string, as taught by Baker, to provide the speaker with essentially visual feedback for quick and easy review of text and to perform revisions (column 4, line 66 – column 5, line 6).

Regarding claims 7 and 14, Keiller in view of Jochumson, Kirby and Brown discloses an apparatus and method for recording speech, to be used as learning data in speech recognition processing, but lacks wherein said presentation means simultaneously displays the recognized character string and the recording character string on a screen by causing unmatched portion or matched portion of at least one recognized character string and the recording character string to blink.

Baker discloses a speech correction device, but does not specifically teach a device wherein said presentation means simultaneously displays the recognized character string and the recording character string on a screen by causing unmatched portion or matched portion of at least one recognized character string and the recording character string to blink.

However, it would have been obvious to one of ordinary skill in the art at the time the invention was made that to provide a visual feedback of the uncertainties by highlighting the instance of uncertainty (e.g. bold or reverse contrast; column 11, lines 22-30 with column 7, lines 1-9) would include flashing, to make the error, mistake and uncertainty stand out.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Keiller in view of Jochumson, Kirby and Brown's apparatus and method wherein said presentation means simultaneously displays the recognized character string and the recording character string on a screen by causing unmatched portion or matched portion of at least one recognized character string and the recording character string to blink, as taught by Baker, to provide the speaker with essentially visual feedback for quick and easy review of text and to perform revisions (column 4, line 66 – column 5, line 6).

Conclusion

7. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Art Unit: 2626

8. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Jakieda R. Jackson whose telephone number is 571-272-7619. The examiner can normally be reached on Monday, Tuesday and Thursday 7:30 a.m. to 5:00p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on 571-272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

JRJ
July 3, 2007


DAVID HUDSPETH
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600